

大規模計算科学研究のこれから

東大院理／東大物性研 常行真司
stsune@phys.s.u-tokyo.ac.jp

スーパーコンピュータの理論ピーク性能は、過去数十年にわたり、5年で10倍程度のペースで向上してきた。このペースは半導体微細加工技術の進展と歩調を一にしたものであり、世界第1位のものでも500位のものでも大差が無い。計算科学は、このようなスーパーコンピュータのハードウェアの進歩に合わせて、研究対象を広げてきた。

計算したい系のサイズ（原子数、電子数、モデルハミルトニアンサイトの数など）を N としたとき、計算量の主要項のオーダーが1回につき N^p であるとしよう。指数 p は計算モデル（電子系、古典粒子系、連続体など）、計算の基礎理論（多体問題の近似手法、計算する物理量など）、計算アルゴリズムによって異なり、通常は1から6程度で、精度が高いほど大きな値になる。時間発展をシミュレーションするなら、計算量はこれに時間発展の長さ T をかけたものである。計算速度が x 年でおおよそ $10^{x/5}$ 倍になることを使って、1日あたりの計算量が必要な計算量に比例するという式をたてると、比例係数を C として $10^{x/5} = C \times N^p \times T$ 、つまり手法が同じなら計算できる系のサイズは $5 \times p$ 年で10倍、時間発展は5年で10倍となり、系が同じであれば計算精度の指標である指数 p の上限は年とともにゆっくり線形に増加する。

計算物質科学の手法開発に携わる研究者は、多体問題を取り扱う**基礎理論**と**アルゴリズム**の工夫で比例係数 C や指数 p を小さくしたり、何らかの**粗視化**や**モデリング**によって系のサイズ N や指数 p を小さくしたり、**探索手法**や**サンプリング**の工夫により同じ結果が得られるまでの時間 T を減らしたりといった努力を続けてきた。またスーパーコンピュータが多数の演算器から構成される並列計算機であることが常識となってからは、理論ピーク性能になるべく近い実効性能を得るための**並列プログラミング**にも、多くの時間と労力を費やしてきた。その結果として、計算機シミュレーションのレベル（サイズ、精度、時間）は、計算機のハードウェアよりも早いペースで向上して来たとし、今後もそうあり続けることを期待する。

一方で、これまでのような**基礎方程式に基づく演繹的シミュレーション**の発展だけに頼るのではなく、既存の小規模シミュレーションを多数の条件（化学組成、初期条件、温度圧力などの環境変数）に対して実行した計算結果から、データ科学の手法を用いて新しい発見につなげる、いわゆる**マテリアルズ・インフォマティクス**も注目されている。これもまた今後の重要な方向性であろう。

本講演では、HPCI戦略プログラム分野2（CMSI）での成果等、実例を紹介しながら、上記のような観点で、物質科学・材料科学における大規模計算科学研究の現状と今後の展開について議論したい。